7.5 / 7.5

Steve Depp
413 – 55

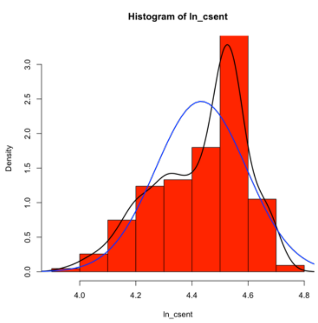
Assignment 4
22 July 2019

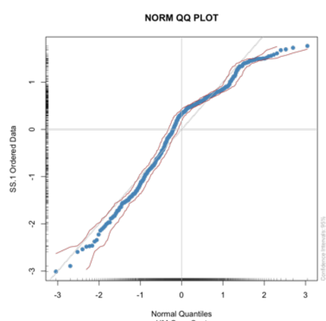
1.1 The following presents sample size, mean, standard deviation, median, 20% trimmed mean, min, max, range, skewness, excess kurtosis for the log of monthly series of Consumer Sentiment of the University of Michigan from January 1, 1978 to August 1, 2013. EDA first confirms there are no gaps in the 428 monthly observations and all dates are 1st of the month. Dates are then decimalized dates now = year + months / 12. The log of original data and not original data is examined here since original data was examined in Assignment 3 and starter code suggests EDA on log data points.

	n	mean	sd	median	trimmed	min	max	range	skew	kurtosis
X1	428	4.432568	0.161774	4.486949	4.442417	3.945458	4.718499	0.7730411	-0.6088029	-0.3681406

The **above** univariates describe a distribution departing from normal via moderate negative skew, and mild negative kurtosis with mild shoulders. Median > mean and a minimum extending more than half the range below center are consistent with negative skew, though trimmed so close to mean would lead one to think the distribution is skewed negatively inside the middle 80% of the distribution.

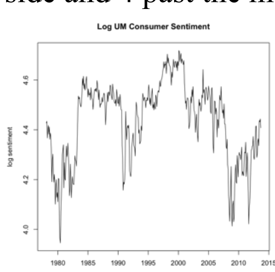
The p-value from a one sample t-test allows us to reject the hypothesis that the distribution is zero means at the highest level of confidence, i.e. > 99.9% as shown on the LHS **below**. A test of null hypotheses that the distribution is symmetric with respect to the mean yields a test statistic = -5.14 and p-value near enough to zero, allowing us to reject the null hypothesis with greater than 99.9% level of confidence. A similar test for kurtosis yields a test statistic -1.55 and p-value 0.12 which does not allow us to reject the null hypothesis of excess kurtosis = 0 with greater than 90% level of confidence.

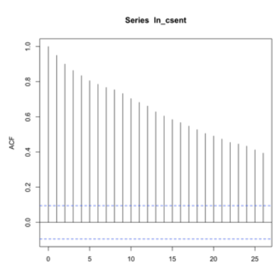


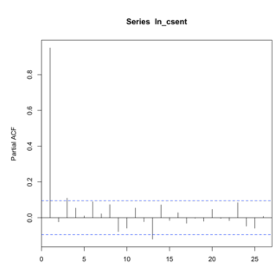



```
        One Sample t-test

data:  ln_csent
t = 566.85, df = 427, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 4.417198 4.447938
sample estimates:
mean of x
 4.432568
```

The density plot **above** has 2 modes, but the histogram reveals just a skewed distribution. Negative kurtosis arising from the prominent left shoulder. The negative skew can be seen in the **above** QQ plot as (a) data minimums extend to 4 SDs on the lower end of the y-axis versus 3 SDs on the left side of the x-axis and (b) data maximums extend only to ~1.7 SDs from mean on the upper end of the y-axis versus 3 SDs on the right side of the x-axis. (This is visible on the density plot with black over blue lines on the LHS.). There are no points past 3 SDs but 2 points extend past the main distribution on the minimum side and 4 past the main distribution on maximum side.







The time series plot **above** shows sharply changing levels, and no trend obvious across the whole range. No consistent change or progression in variance is evident across the range though variance does vary in the data and in ranges of the data. No discernable seasonality from this distance. Some discernable pattern to cyclicality but nothing to depend upon: 7 years sideways with in ranges then shifts sharply lower. Progressions higher are more trend like, but with initial variation that would render them unpredictable. In general, these patterns are aperiodic and unreliable.

The ACF plot **in center above** with steadily declining autocorrelation levels is consistent with the single near unity partial correlation at lag = 1 in the PACF plot on the **RHS above** and suggests that a 1st difference would likely remove autocorrelations from both plots. The significantly negative partial autocorrelation at lag = 13 should be investigated for whether its appearance is consistent with the randomness of a 95% confidence interval or truly represents some form of 13-month seasonality.

```
Title:
 Augmented Dickey-Fuller Test

Test Results:
  PARAMETER:
    Lag Order: 2
  STATISTIC:
    Dickey-Fuller: -2.9398
  P VALUE:
    0.04357
```

```
Title:
 Augmented Dickey-Fuller Test

Test Results:
  PARAMETER:
    Lag Order: 7
  STATISTIC:
    Dickey-Fuller: -2.1627
  P VALUE:
    0.252
```

```
Title:
 Augmented Dickey-Fuller Test

Test Results:
  PARAMETER:
    Lag Order: 10
  STATISTIC:
    Dickey-Fuller: -2.3068
  P VALUE:
    0.1983
```

```
Title:
 Augmented Dickey-Fuller Test

Test Results:
  PARAMETER:
    Lag Order: 13
  STATISTIC:
    Dickey-Fuller: -2.423
  P VALUE:
    0.1549
```

There are several tests for unit roots to apply to this data. **Firstly**, the Augmented Dickey Fuller Test's null hypothesis is that there are unit roots and the data is not stationary. The alternate hypotheses are that the series is stationary or trend stationary with a constant but no trend (type = "c") that appears to fit our raw data. Testing the null hypothesis with a lag = 2, 7, 10, and 13 months, **above**, we are unable to reject the null hypothesis that our data is non-stationary and has unit roots. 7, 10, and 13 months are the local and absolute (lag=7) maxima of p-values inside lag=16. (Lag 2 may suggest less structure at 1 and 2 lags, but I need to read more on this.) **Secondly**, KPSS Unit Root Test is another test which automatically employs lag = 5 for our data set and tests the null hypothesis that our data are stationary. Here **below** the results yield a test statistic 0.8793 > critical value = 0.739 which allows us to reject the null hypothesis that the data is stationary at the 99% confidence level. **Finally**, KPSS tests can be repeatedly executed via the 'ndiffs' function to determine the appropriate number of first differences, which in this case is 1. This appears to betray the lowest p-value obtained from the Dickey Fuller test at lag = 2.

```
#######################
# KPSS Unit Root Test #
#######################

Test is of type: mu with 5 lags.

Value of test-statistic is: 0.8793

Critical value for a significance level of:
                10pct  5pct 2.5pct  1pct
critical values 0.347 0.463  0.574 0.739
```

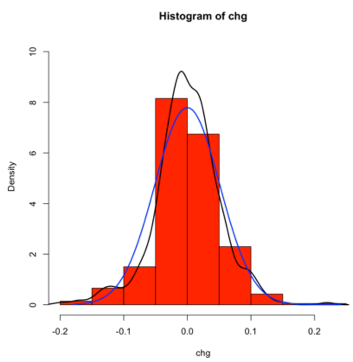
1.2 The results of the above tests suggest there are unit roots to this data which objectively determine whether differencing is required in order to make the data stationary. In this case it appears 1st differencing is warranted, looking at the results of the 'ndiffs' test, the PACF at lag=1 and the gradual but scalloped decline in ACF lines.

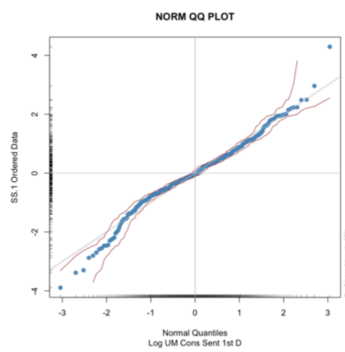
1.3 The table **below** presents sample size, mean, standard deviation, median, 20% trimmed mean, min, max, range, skewness, excess kurtosis for the 1st differences of the log of monthly Consumer Sentiment from January 1, 1978 to August 1, 2013.

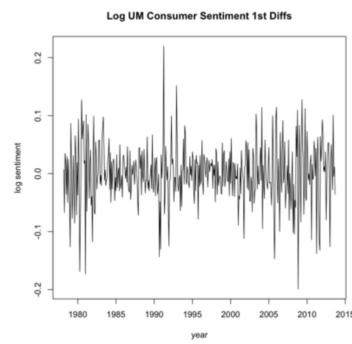
	n	mean	sd	median	trimmed	min	max	range	skew	kurtosis
X1	427	-4.520131e-05	0.05118744	-0.002162163	0.0008734556	-0.1992492	0.2197286	0.4189779	-0.1995997	1.730398

```
        One Sample t-test

data:  chg
t = -0.018247, df = 426, p-value = 0.9855
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.004914127  0.004823725
sample estimates:
    mean of x
-4.520131e-05
```

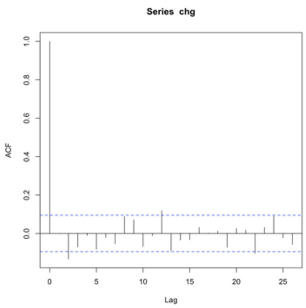
First differences' mean = 0 = median = trimmed mean and all lie in the middle of the range which is consistent with very mild skew = -0.20. The null hypotheses of true mean = 0 cannot be rejected at all (as shown **above**) and the null hypothesis of an unskewed distribution can be rejected only at the 90% level, but cannot be rejected with any higher level of confidence as suggested by the test statistic = -1.68 and associated p-value = 0.0922. On the other hand, a null hypothesis of no excess kurtosis can be rejected with a test statistic = 7.29 with greater than 99.9% confidence. Excess kurtosis = 1.73.

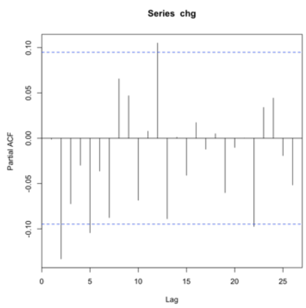






It would appear from the **above** plots the distribution rapidly approaches normal after 1st differencing logs. One can slight negative skew of the mode in the histogram and peakedness and thin shoulders relative to normal in the density plots, consistent with excess kurtosis. Similarly, QQ plot also **above** shows most data points lie within -3SDs to +2.5SD, consistent with moderate skew, and an uneven distribution of outliers beyond 3SDs on both sides: only 1 stray at the upper end of our range versus 3 at the lower end. The points in our data at -3SDs would normally fall at 2.5SDs again consistent with moderate negative skew; this is more prominent on the LHS than RHS. The time series of the 1st difference of logs, **above**, appears stationary around a constant mean with alternating but no patterned level of variance. One can discern more narrow ranges +/- 0.05 than might normally be seen, which might be consistent with peakedness of the fat tailed distribution. Net, the 1st differences distribution is more suitable for modeling.






```
Title:
 Augmented Dickey-Fuller Test

Test Results:
  PARAMETER:
    Lag Order: 10
  STATISTIC:
    Dickey-Fuller: -7.0281
  P VALUE:
    0.01
```

```
#######################
# KPSS Unit Root Test #
#######################

Test is of type: mu with 5 lags.

Value of test-statistic is: 0.0331

Critical value for a significance level of:
                10pct  5pct 2.5pct  1pct
critical values 0.347 0.463  0.574 0.739
```

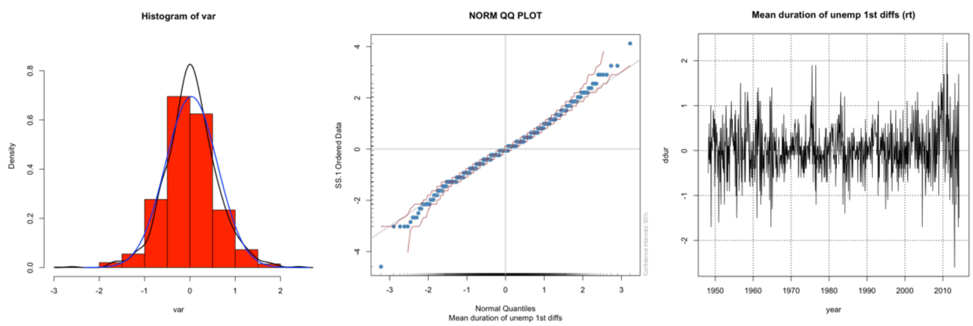
A distribution consistent with stationary is also represented by the **above** ACF and PACF plots where any of the near significant lag correlations might be due to random sampling than a serially correlated 1st differenced log observations. The Dickey Fuller test at all lags produces p-values enabling us to reject the null hypothesis of a unit root. With KPSS Unit Root test statistic = 0.0331, one cannot reject the null hypothesis of stationarity.

1.4 The differenced log data is stationary with constant mean and seemingly constant variance with no observable seasonality where the undifferenced log data was not stationary, subject to changing means and variances across different sub segments of the range.

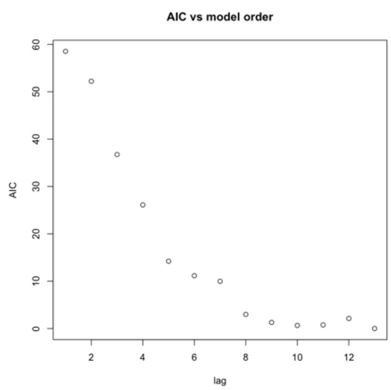
```
        One Sample t-test

data:  ddur
t = 1.6507, df = 793, p-value = 0.0992
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.006362152  0.073616560
sample estimates:
mean of x
0.0336272
```

	n	mean	sd	median	trimmed	min	max	range	skew	kurtosis
X1	794	0.0336272	0.5740414	0	0.02735849	-2.6	2.4	5	0.04897172	1.255041

2.1 The one sample t-test statistic, shown **above**, yields a 0.0992 p-value and associated 95% confidence interval stretching from -0.00636 to 0.07362. The confidence interval would enable us to reject the null hypothesis of mean = 0 with 95% confidence, but the p-value would prevents us from rejecting the null hypothesis of zero mean with 95% confidence on the basis of the p-value, but would allow us to reject the null hypothesis with 90% confidence. The p-value is centered on zero mean as the true population mean whereas the confidence interval surrounds the sample mean and just misses zero. Half the confidence interval is ~0.0368 which if it were the standard error would explain why a sample mean of 0.03363 as we observe would be nearing significant. The true test however is stationed over the true hypothesized mean of zero and thus, we cannot reject the hypothesis with any more than 90% confidence, and just barely that. Any conclusion we make is predicated upon a distribution with true mean = zero and the distribution from which standard errors can be compared with a z-score from an equivalent normal distribution. In this case, so close to 90% and so far from 95% confidence, the appropriate steps would be to examine the distribution for skew, excess kurtosis, normality, constant variance, and serial correlation. In the above **table**, our distribution shows excess kurtosis, moderate skew, a mean in the northern half of the range near its median and below the trimmed mean.

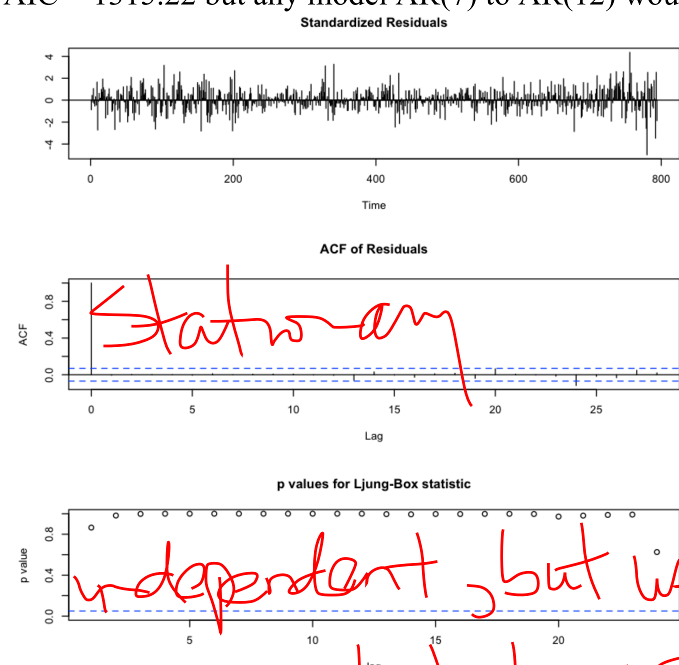


A very quick review of plots of density/histogram, QQ, and time series does not betray that the mean is indeed near 0 and the distribution approaching normal.



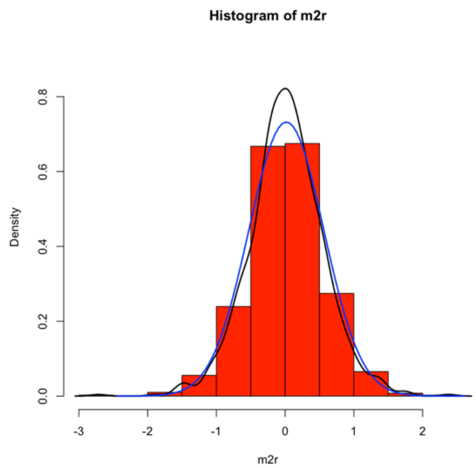
0	58.526
1	52.216
2	36.741
3	26.102
4	14.213
5	11.156
6	9.979
7	2.986
8	1.292
9	0.648
10	0.766
11	2.13
12	0

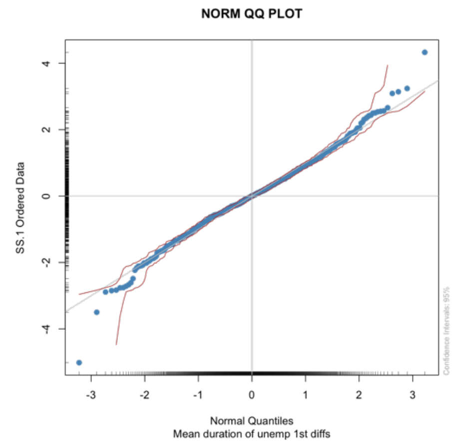
2.2 An examination of model order between AR(1) to AR(12) on the basis of incremental AIC yields the table **above**, with order in bold and incremental AIC to the right of it, and plot of order vs incremental AIC. On this basis, one would select an AR(12) model whose AIC = 1315.22 but any model AR(7) to AR(12) would suffice on an AIC basis.



Stationary

independent, but use Ljung-Box test as plot biased





constant

A diagnostic review of AR(12) model residuals **above** top presents apparent zero mean constant and un predictable variance with a degree of expanding, alternating variance in the last ~50 observations. The ACF and p-values for Ljung-Box test present a pleasant view of these models. Box-Ljung test have a p-values nearing 100% for models of order 1 to 23 indicating we cannot reject the hypothesis of zero correlations from lag = 1 to 24. The density and QQ plots appear to be normal across the time range plotted. 1

	n	mean	sd	median	trimmed	min	max	range	skew	kurtosis
X1	794	0.01766675	0.5448237	0.01861794	0.01787306	-2.715031	2.37558	5.090611	-0.05611392	1.435885

Residuals univariate measures in the table **above** show moderate negative skew, excess kurtosis, and a median and trimmed mean above the mean.

```
            One Sample t-test

data:  m2r
t = 0.91372, df = 793, p-value = 0.3611
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.02028721  0.05562072
sample estimates:
 mean of x
0.01766675
```

The t-test pvalue = 0.3611 and the 95% confidence interval **above** suggests the mean is zero. Test statistics -0.64 for skew and 8.25 for excess kurtosis confirm that we cannot reject the hypothesis of non-zero skew but can reject the hypothesis of non-zero excess kurtosis.

	n	mean	sd	median	trimmed	min	max	range	skew	kurtosis
X1	397	0.008791397	0.5305801	0.00380354	0.005900746	-1.533922	1.782508	3.31643	0.09393427	0.6933406

	n	mean	sd	median	trimmed	min	max	range	skew	kurtosis
X1	397	0.02654211	0.5592334	0.0317905	0.0302038	-2.715031	2.37558	5.090611	-0.1887389	2.013709

Dividing the data in left half = early data on the time series and right half = more recent time stamps, the means are different but still lie within the 95% confidence interval for the whole data set. Separate t-tests confirm both are likely zero with p-values of 0.7415 for the early data and 0.3449 for the latter data. Standard deviations are similar as well. So I believe we can suggest these residuals are stationary and the model is adequate.

2.3 The equation for our fitted model AR(12) is

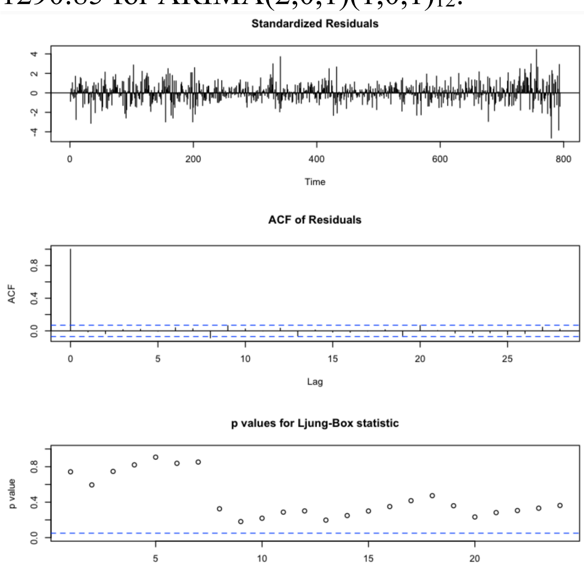
$$(1 - phi_1{*}B^1 - phi_2{*}B^2 - phi_3{*}B^3 - phi_4{*}B^4 - phi_5{*}B^5 - phi_6{*}B^6 \\ - phi_7{*}B^7 - phi_8{*}B^8 - phi_9{*}B^9 - phi_{10}{*}B^{10} - phi_{11}{*}B^{11} - phi_{12}{*}B^{12})\ (1 - B^1)\ \ln(r_t)$$

ar1	-0.1351
ar2	0.1134
ar3	0.1083
ar4	0.1165
ar5	0.0738
ar6	0.077
ar7	0.0903
ar8	-0.0609
ar9	0.0687
ar10	0.0609
ar11	0.0228
ar12	-0.0723

Where the coefficients phi_i are identified here as 'ar1' to 'ar12'**above.**

```
Call:
arima(x = ddur, order = c(2, 0, 1), seasonal = list(order = c(1, 0, 1), period = 12),
    include.mean = F)

Coefficients:
         ar1     ar2      ma1    sar1     sma1
      0.6538  0.2638  -0.8021  0.5662  -0.7429
s.e.  0.0478  0.0360   0.0382  0.0755   0.0585

sigma^2 estimated as 0.2926:  log likelihood = -639.43,  aic = 1290.85
```

2.4 The seasonal model shown above achieves an AIC reduction from 1315.22 in AR(12) to 1290.85 for ARIMA(2,0,1)(1,0,1)$_{12}$.



as above

Diagnostic plots **above** are representative of what appears to be stationary residuals with mean zero and constant or at least random variance and Box Ljung test suggesting that portmanteau residual correlations are confidently zero from lag 1 to lag 8. Lag 9, 13, 20 may approach p-values where correlations are zero with only 90% confidence.

	n	mean	sd	median	trimmed	min	max	range	skew	kurtosis
X1	794	0.02208124	0.5408249	0.01980199	0.02173899	-2.513378	2.412286	4.925664	-0.03507631	1.577368

```
        One Sample t-test

data:  m4r
t = 1.1505, df = 793, p-value = 0.2503
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.01559416  0.05975664
sample estimates:
 mean of x
0.02208124
```

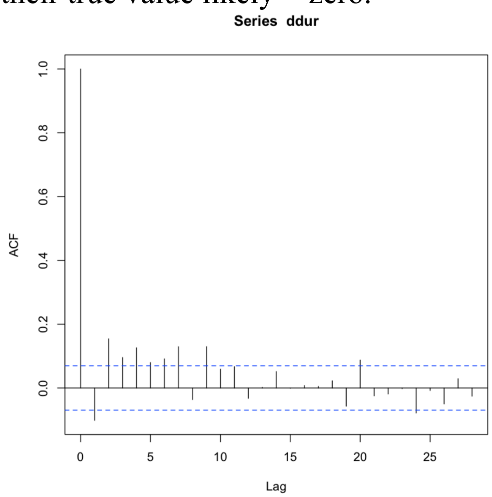
The above tables shows mean, median and trimmed mean near to each other, skew that is moderate and kurtosis which is elevated. One sample t-test confirms the zero mean, via a 0.2503 p-value preventing us from rejecting the null hypothesis of zero mean. Similar tests of skew and kurtosis, confirm skew is likely zero and kurtosis non zero.

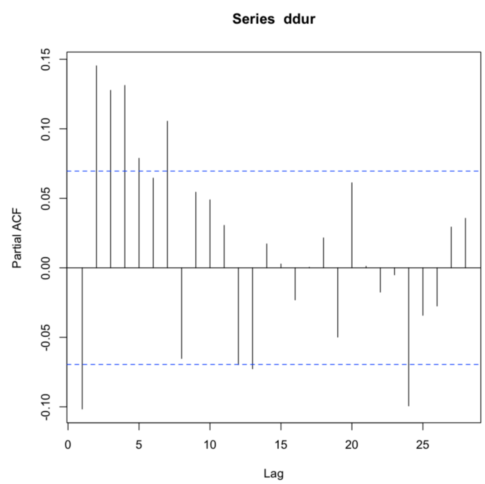
	n	mean	sd	median	trimmed	min	max	range	skew	kurtosis
X1	264	-0.0003568065	0.5672601	0.004024129	0.005218567	-1.693341	1.550373	3.243714	-0.1610301	0.3060764

	n	mean	sd	median	trimmed	min	max	range	skew	kurtosis
X1	264	0.02220614	0.4295	0.007989941	0.01442544	-1.252966	2.010643	3.263609	0.5238817	2.299825

	n	mean	sd	median	trimmed	min	max	range	skew	kurtosis
X1	266	0.04422661	0.6097868	0.04732059	0.04647524	-2.513378	2.412286	4.925664	-0.1404549	1.688787

Dividing the distribution up into earliest (**top above**), middle dates (**middle above**), and most recent time series data points, one can see the differences in distribution, notably narrower standard deviations in the middle dates and wider ranges in the more recent third of the data. The means for these residuals test with pvalues 0.24 to 0.99 confirming their true value likely = zero.



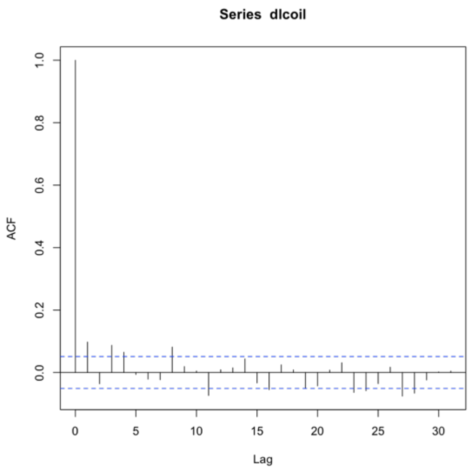


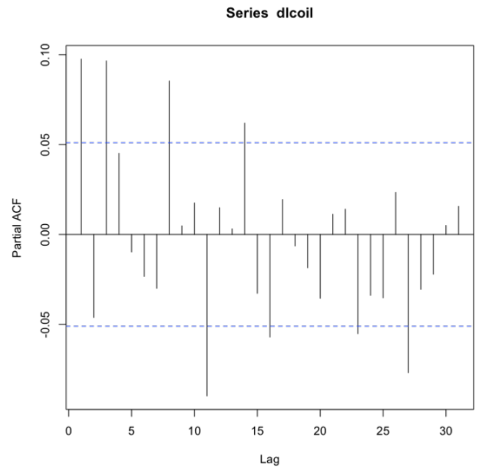
2.5 Based upon the models presented thus far and their AIC levels, the latter seasonal model achieves lower AIC levels and captures the lag 12 component of the AR(12) model, if that is even appropriate. Looking at the the ACF and PACF plots of ddur, however, one questions whether this is more appropriately a MA model entirely. The absolute magnitude of ACF exceeds critical values for lags 1-7 and 9, and then drops off either at 8 or at 10. The absolute magnitude of PACF decline steadily (unsure exponentially) to either lag 7 or 8 and find some resilience at lag 12 and 13. Taken together, these would be consistent with MA via seasons when employing the highlighted cells of the Rules of Thumb **below**:

Table 1: Rules of Thumb.

Function	MA(q)	AR(p)	ARMA(p,q)	White Noise
ACF	D(q)	T	T	0
PACF	T	D(p)	T	0
IACF	T	D(p)	T	0

2.6 Employing the backtest of both models, one would prefer the seasonal model to the AR(12) model due to reduced RMSE of the former. The seasonal model RMSE = 0.944 where the AR(12) yields 0.972 RMSE.






```
        Box-Ljung test

data:  rtn
X-squared = 45.533, df = 10, p-value = 1.742e-06
```

3.1 There appears to be serial correlation of the 1st difference of log oil prices, (dlcoil), in the ACF and PACF plots **above**. In particular, in ACF at lag = 1, 3, 4, 7 and 11 and in PACF at 1, 3, 8 and 11. Employing Box-Ljung, we observe a p-value near zero suggesting we can reject the null hypothesis of zero correlation out to lag = 10.

```
Call:
arima(x = rtn, order = c(11, 0, 0))

Coefficients:
         ar1      ar2     ar3     ar4      ar5      ar6      ar7     ar8
      0.1055  -0.0513  0.1054  0.0421  -0.0192  -0.0160  -0.0381  0.0980
s.e.  0.0259   0.0261  0.0263  0.0263   0.0264   0.0264   0.0265  0.0265
          ar9    ar10     ar11  intercept
      -0.0041  0.0303  -0.0942     0.0009
s.e.   0.0266  0.0266   0.0264     0.0013

sigma^2 estimated as 0.001827:  log likelihood = 2553.47,  aic = -5080.93
```

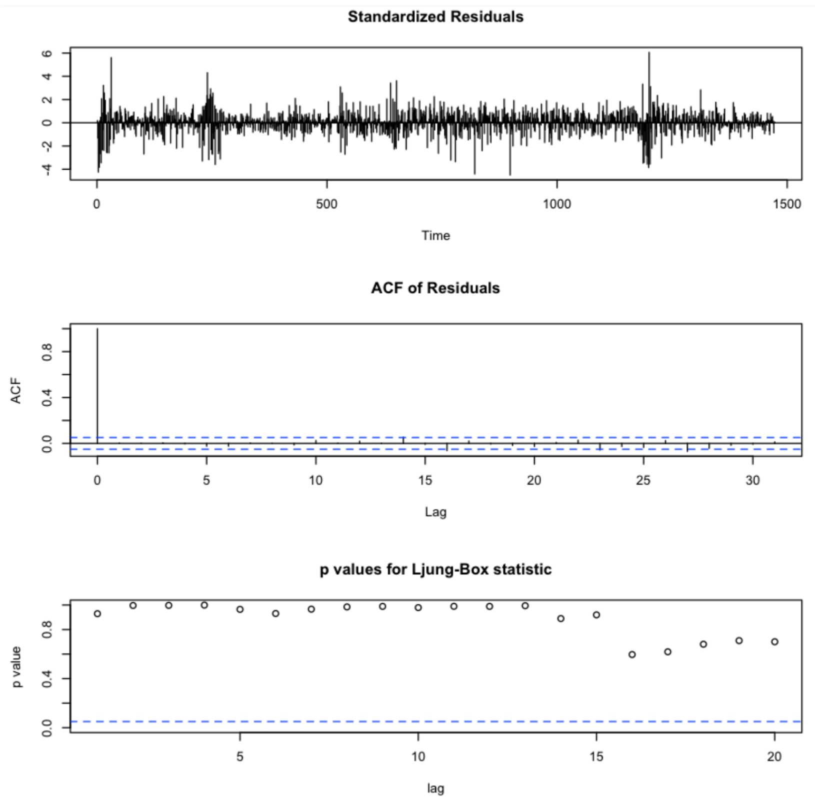
3.2 Employing the model code yields an AR(11) model **above** with AIC = -5080.93 and then removing ar5, ar6, ar9 and ar10 whose values fall short of their standard error, we get:

```
Call:
arima(x = rtn, order = c(11, 0, 0), include.mean = F, fixed = c1)

Coefficients:
         ar1      ar2     ar3     ar4  ar5  ar6      ar7     ar8  ar9  ar10
      0.1050  -0.0510  0.1044  0.0408    0    0  -0.0349  0.0949    0     0
s.e.  0.0258   0.0258  0.0261  0.0261    0    0   0.0261  0.0261    0     0
         ar11
      -0.0897
s.e.   0.0262

sigma^2 estimated as 0.00183:  log likelihood = 2552.09,  aic = -5088.19
```

Zeroing these coefficients produces a simpler model with -5088.19 AIC.
The equation for this model would be:
$(1 - phi_1 * B^1 - phi_2 * B^2 - phi_3 * B^3 - phi_4 * B^4 - phi_7 * B^7 - phi_8 * B^8 - phi_{11} * B^{11})) (1 - B^1) \ln(r_t)$

With actual coefficients:

$$(1 - 0.105{*}B^1 - (-0.051){*}B^2 - 0.1044{*}B^3 - 0.0408{*}B^4 - (-0.0349){*}B^7 - 0.0949{*}B^8 - (-0.0897){*}B^{11})\ (1 - B^1)\ \ln(r_t)$$

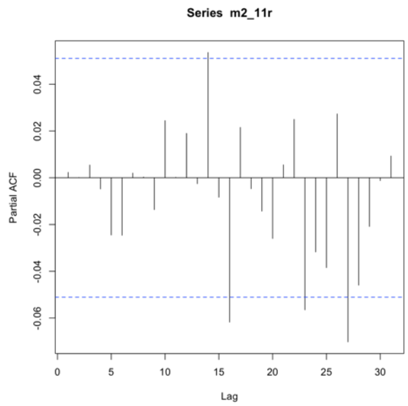


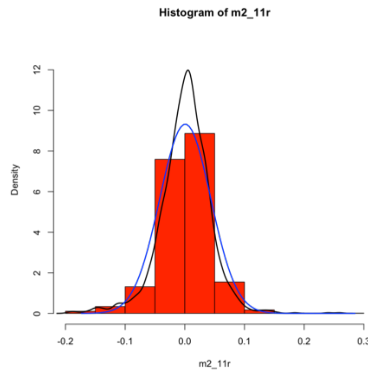
From these plots of residuals, ACF, and Ljung-Box p-values, it appears the residuals for the above model are uncorrelated as a group out to lag = 15, carry no significant serial correlation and seem distributed evenly about a zero mean with constant variance.

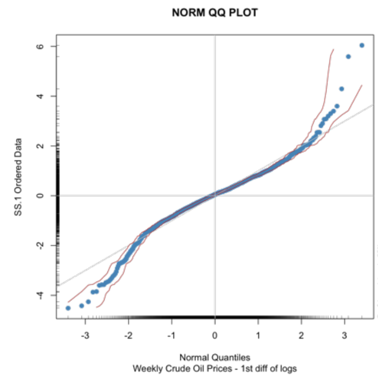
```
        One Sample t-test

data:  m2r
t = 0.68454, df = 1472, p-value = 0.4937
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.001423842  0.002950301
sample estimates:
   mean of x
0.0007632292
```

	n	mean	sd	median	trimmed	min	max	range	skew	kurtosis
X1	1473	0.0007632292	0.04279165	0.002948072	0.001859475	-0.1925152	0.2593203	0.4518355	-0.1994634	3.441333

Univariates confirm zero mean=median=trimmed mean as does the one sample t-test **above**. Skew is very mildly negative, excessive kurtosis high but both are confirmed by their p-value to be non-zero.



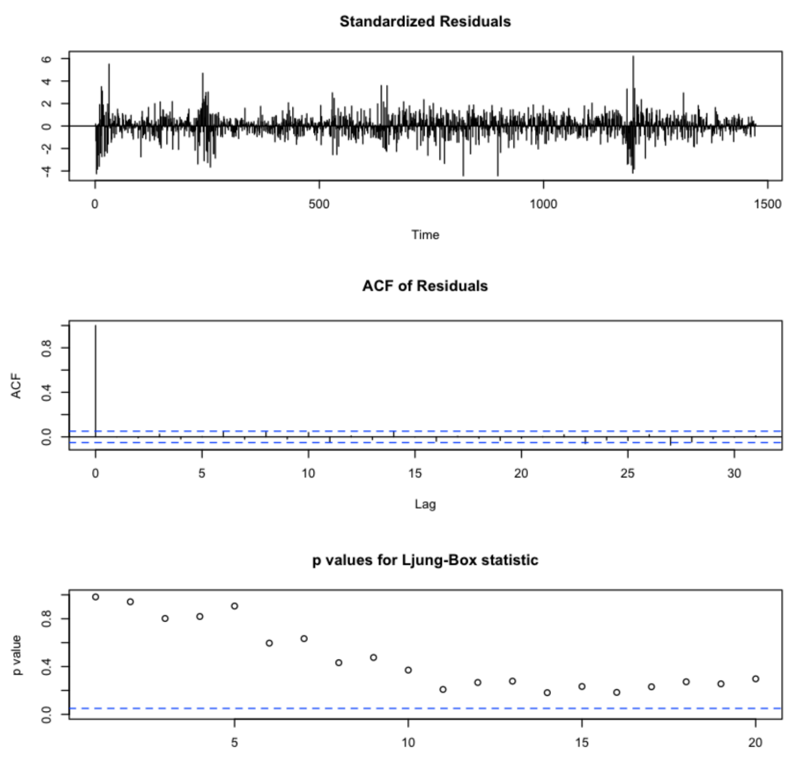




Partial ACF of residuals confirms that autocorrelation to lag = 14 are not significant, but beyond lag=15 there may be structure worth modeling. While histograms and density of residuals appears normal, the QQ plot exhibits outliers which should be investigated for influence. There are 3 outliers on the RHS and 5 on the LHS and the LHS has a longer tail that might skew the model.

```
Call:
arima(x = rtn, order = c(3, 0, 2), include.mean = F)

Coefficients:
         ar1      ar2     ar3      ma1     ma2
      0.5664  -0.8548  0.1689  -0.4680  0.7753
s.e.  0.0934   0.0681  0.0270   0.0931  0.0680

sigma^2 estimated as 0.001845:  log likelihood = 2546.4,  aic = -5080.8
```

3.3 Fitting a final model ARMA (3,0,2) to the 1st differenced log oil price data yields the **above** coefficients and AIC 5080.80 which is close to the original AR(11) model before deleting a few parameters. This model is simpler than the other 2 and so would be a candidate for selection.

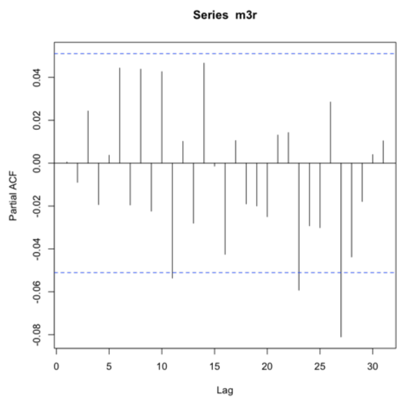


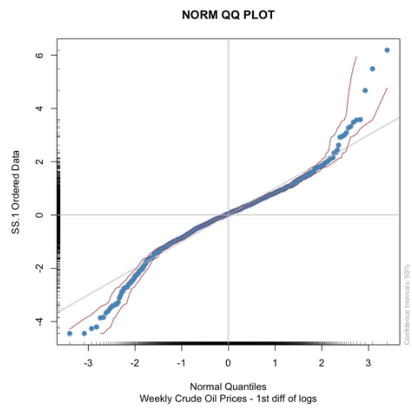
From residuals **above**, it would appear that serial correlation is removed, mean is zero'd out and variance is somewhat consistent across all time segments. Box-Ljung out to lag 11 is confidently zero.

	n	mean	sd	median	trimmed	min	max	range	skew	kurtosis
X1	1473	0.0007899324	0.04295793	0.002585217	0.001823346	-0.19032	0.2666097	0.4569297	-0.1566546	3.700366

```
        One Sample t-test

data:  m3r
t = 0.70575, df = 1472, p-value = 0.4805
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -0.001405638  0.002985502
sample estimates:
   mean of x
0.0007899324
```

A review of the residuals univariates and t-test of the mean **above**, the distribution is zero meaned with slight negative skew and high excess kurtosis.





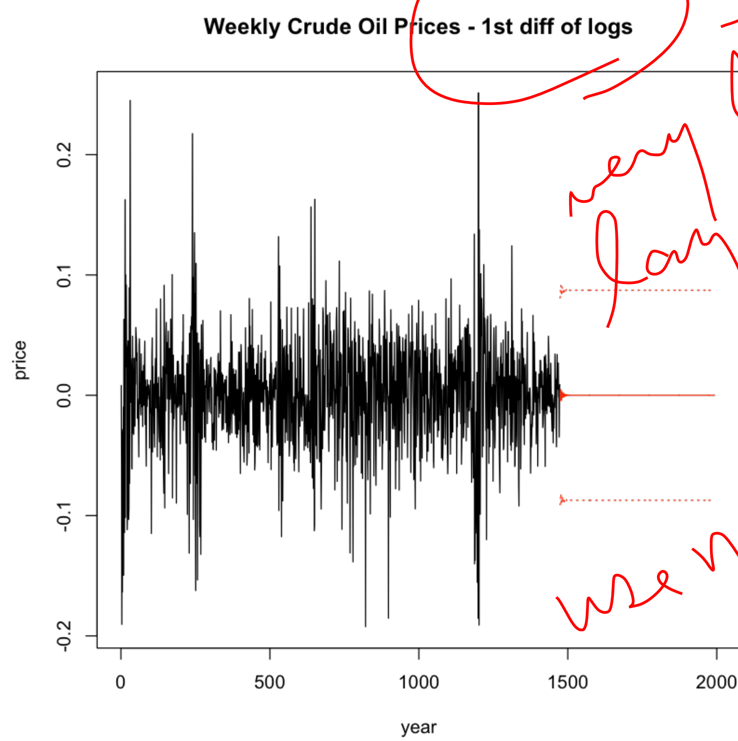
Residuals PACF and QQ are similar to the previous model with no significant autocorrelation and very kurtotic / fat tailed distribution which should be investigated. Otherwise the model appears to be adequate.

The formula for this model would be

$(1 - \text{phi}_1{*}B^1 - \text{phi}_2{*}B^2 - \text{phi}_3{*}B^3)\ (1 - B^1)\ \ln(r_t) = (1 - \text{theta}_1{*}B^1 - \text{theta}_2{*}B^2)\ (\text{alpha}_t)$

$(1 - 0.5664{*}B^1 - (-0.8548){*}B^2 - 0.1689{*}B^3)\ (1 - B^1)\ \ln(r_t)$
$= (1 - (-0.468){*}B^1 - 0.7753{*}B^2)\ (\text{alpha}_t)$

3.4 Solely on the basis of fit, I would not select one model over the other, but again, the feel of the ACF and PACF suggests that some degree of residual modeling is required and thus MA must be included. I would select the last model as a result despite the lower absolute value AIC -5080.8 which is marginally worse than the amended AR(11) model = -5088.
The AR(11) model performs better on an RMSE basis 0.9683 than the ARMA(3,2) which scored 0.9916. Looking at MAE though, the 2 models are quite similar, 0.7794 and 0.8002 respectively. Since we are dealing with fat tailed residuals, I am more inclined to lean on MAE and remain undecided on this basis, sticking with a decision to employ the ARMA model.



this means very little to lay persons

use natural scale

4.0 The selected model extracts information from changes in weekly oil prices, looking back 1, 2 and 3 weeks at the log of oil prices and then back 1 and 2 weeks at the errors of modeling oil prices. All this sounds complex, but this is the best model for forecasting oil into the future. We feel this is the case because we looked back at the last 15 years of data and found this model performed equally well as less complicated models, and because it allows us to look forward at the bands of possible oil price changes in the future while accounting for model errors along the way. As for action, were the daily difference in log oil prices to exceed these bounds, you could sell at the upper bound and buy at the lower bound and likely make money 95% of the time.